Object Recognition of Environmental Information in the Internet of Things Based on Augmented Reality

Chen Shasha1, Wang Mei1*, and Qin Xuebin1

1Xi’an University of Science and Technology, Xi’an, Shaanx, China

Abstract. With the rapid development of Internet of things, on the basis of existing mine technology, as the basic digital mine, the detection system will become more and more mature with the construction of digital mines. This article takes China's coal production as the main basis, developed a matching coal mine augmented reality system. This article uses augmented reality technology to provide information through computer systems, increase the perception of coal mine workers on the real environment, and superimpose computer-generated prompt information into the real scene of the coal mine. Realize "enhancement" of real coal mine environment. The experimental results show that this system is workable. Object recognition of environmental information in the Internet of things based on augmented reality system designed and implemented in this paper, in providing a new direction for Research on IoT Smart Mines and Application of Mobile Devices.

1 Introduction

The rapid progress of information and network technology has evolved from the Internet and automation systems that were originally used for administrative offices and industrial and commercial applications to the application of these technologies everywhere in life. The Internet has also become increasingly popular. Every household has network coverage. People began to pursue a more convenient and better living environment, and began to study the application of mobile devices, apps, and mobile networks in environmental monitoring, machine automation, smart home, etc. These applications will use the Internet of Things (Internet of Things). IoT, and augmented reality technology has provided new opportunities for the popularization of the Internet of Things. On January 26, 2015, Microsoft released HoloLens, an augmented reality technology device for holographic glasses. HoloLens is a device that displays holographic images in real-life situations. These virtual holographic images can be used by the user to interact with remote devices and environments with gesture control, which is very beneficial to the future development of the Internet of Things. Research and research on patent analysis combined with Internet of Things and augmented reality technology. Taking Microsoft HoloLens as an example, look from the patent layout, look development from the platform, analyze the development trend of future personal intelligent terminal equipment, and provide a path for China’s future development of the Internet of Things. make a suggestion.

Augmented reality is a technology that fuses computer-generated virtual objects with other information and the real environment the user sees. It has the features of virtual reality, real-time interaction, and three-digit registration. Compared with traditional virtual reality (VR), the biggest difference between augmented reality is that augmented reality does not require a realistic and complete virtual world. It only needs to display part of the information in the virtual world in order to assist the real world. This also reduces the requirements for computing power. Compared with immersive virtual reality, augmented reality is not only widely used but also more secure. Because the augmented reality system allows users to see the virtual object, they can still see the real scene.

2 Augmented Reality System and Its Main Theories

Augmented reality is a technique that calculates the position and angle of a camera image in real time and adds a corresponding image. It is generally believed that the emergence of augmented reality technology stems from the development of virtual reality technology, but there are obvious differences between the two. Traditional virtual reality technology gives users a completely immersive experience in the virtual world, creating another world. Augmented reality is a new technology that seamlessly integrates information in the real world and information in the virtual world. It integrates physical information (visual information, sound, taste, touch, etc.) that was originally difficult to experience in the real world for a certain period of time and space through computer science and technology, simulates and overlays it, and applies virtual information...
to the real world. Sensed by human senses to achieve a sensory experience that transcends reality. The real environment and virtual objects are superimposed on the same scene or space in real time.

![Fig. 2.1 Structure of Augmented Reality System.](image)

### 2.1 Tracking Registration

Tracking registration refers to digitizing a target object by calculating it, using a camera of a smart device and the relative position of a captured target object to identify the target object. Through calculation, a projection matrix is obtained. The virtual information is passed through these projection matrices. To instantiate, at the same time, the virtual information is correctly delivered to the correct location to form the final image we need. The key to this system is the registration technology. The advantages and disadvantages of the registration technology are directly related to the success of the AR system. In the current popular AR system, according to the difference of the tracking principle, the tracking registration technology can generally be divided into three categories: tracking registration technology based on hardware sensor devices, tracking registration technology based on computer vision, and hybrid tracking technology, as the picture shows.

![Fig. 2.2 Three-dimensional registration technology classification.](image)

Marker is also included in the system, which is also called strong mark-based tracking registration technology. Marker's main role is to enable the system to quickly detect the existence of Marker in a complex and chaotic reality environment, so that the virtual information can be registered within the space of Marker's location.

The simplest Marker is made up of black rectangles and various graphics or text within the rectangle. As shown below, Marker's pattern contains different virtual information, so the method of extracting Marker is not the same. Therefore, a reasonable Marker should be selected to improve the accuracy of recognition.

A set of video frames collected by the camera is matched in real time to determine the matching feature point gamma between the virtual historical resource information and the real object carrier of the cultural object, so as to obtain a single mapping relationship between the two. In the actual registration process, the transformation between the target coordinate system, the image plane coordinate system, and the marker coordinate system is involved. The specific transformation relationship is shown in Figure 2.

![Fig. 2.3 Relationship between camera and tracking target.](image)

\[
\begin{bmatrix}
X_c \\
Y_c \\
Z_c
\end{bmatrix} = [V_1, V_2, V_3 | W] \begin{bmatrix}
X_b \\
Y_b \\
Z_b
\end{bmatrix} = (V_1 \ast V_2)
\]

The above formula is a coordinate conversion equation between the camera and the plane marker.

### 2.2 Augmented Reality Display

Microsoft HoloLens, a special eyewear that incorporates CPU, GPU, and holographic processors, allows users to enter the full virtual world at home, using images, sounds, and images to access the holographic experience using the surrounding environment as a vehicle. The headset includes a transparent display on the black lenses, and the stereo system not only allows the user to see, but also hears sounds from surrounding holograms. At the same time, HoloLens also has a built-in set of sensors to implement each Functionality. Hololens3 features:

1. Looks like VR goggles, in fact, is AR head display: HoloLens has a vision penetrating all-transmission mirror, it can use advanced optical projection system to generate multi-dimensional full-color holographic images, and has a very low delay, so it can let the user get a holographic experience, the key point is that there are enough spots in the holographic image, that is to have a high holographic density, while fixing these points to the world around the user. HoloLens has a holographic density of over 2500 spots per arc, so many spots provide the user a richer, brighter, and more full holographic image. HoloLens has built-in advanced sensors for acquiring user behavior and environment. Its built-in 4 environment-aware cameras, a depth camera, and a 2 megapixel Full HD video camera can also be used for recording. We call it Mixed Reality Capture (MRC), which allows users to Others share high-definition video with holograms.

2. This is a PC! Microsoft's HoloLens is by far the first fully wireless Windows 10 device. It is also equipped with Microsoft's own Windows Holographic dedicated OS, CPU, Memory, Storage, and peripheral cameras. The accelerometer is also equipped with a body.
He himself includes speakers, power supply, Wi-fi 802.11ac, micro-USB 2.0 interface, so he does not need to connect to the PC by cable and telephone. Unlike CPUs and GPUs, HoloLens uses Microsoft's custom 32-bit Holographic Processing Unit (HPU). This custom processor efficiently handles the large amount of data that the sensor sends every second.

3. Application of AR goggles: Through HoloLens, users can draw holograms with gestures and interact with applications with voice and line of sight. A spatial sound field that can be realized for synthesizing sound anywhere in the space, even behind the user. At the same time, it can achieve the most natural human-computer interaction modes such as eye tracking, gesture input, and voice support. HoloLens can sense the user's line of sight and present the world around the user in real time. For entertainment content such as games and videos, Microsoft HoloLens focuses on 2B applications: aerospace, medical training, construction, and manufacturing equipment maintenance.

![Fig. 2.4 A partial exploded view of Microsoft HoloLens.](image)

### 2.3 AR system software based on unity3D

The Unity3D game engine developed by Unity Technologies, Denmark, provides a platform for players around the world to easily build 3D video games, visual interactive building displays, and 3D animations in real-time interactive multi-platform comprehensive large-scale game development tools, which can be called comprehensive, professional Game Engine.

![Fig. 2.5 Unity3D function modules.](image)

The main features of Unity3D are:

1. Visual programming interface completes various development work, efficient script editing, convenient development;
2. Automatic instant import, Unity supports most 3D models, bones and animations are imported directly, and the texture material is automatically converted to U3D format. According to its excellent real-time 3D graphics rendering technology, it can excellently complete the display of virtual information, enabling developers to While developing

3. Multi-platform development and deployment of the work can be accomplished with just one click. The most powerful feature of Unity3D is its cross-platform publishing features.
4. the bottom layer supports OpenGL and Direct11, simple and practical physics engine, high-quality particle system, easy to use, the effect is realistic;
5. multi-script language development. Support Java Script, C#, Boo scripting language, easy to use, high development efficiency;
6. Unity has excellent performance, outstanding development efficiency, and great cost-effectiveness;
7. support from stand-alone applications to large multiplayer networked game development.

<table>
<thead>
<tr>
<th>Name</th>
<th>Publishing platform</th>
<th>language</th>
<th>Open Source</th>
<th>efficiency</th>
<th>environment</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cry Engine</td>
<td>Win,X64,0,PS3</td>
<td>C++,Visual Script</td>
<td>N</td>
<td>L</td>
<td>T</td>
<td>H</td>
</tr>
<tr>
<td>Unreal Engine</td>
<td>Win,X64,0,PS3</td>
<td>C++,Visual Script</td>
<td>Y</td>
<td>L</td>
<td>T</td>
<td>L</td>
</tr>
<tr>
<td>Unity 3D</td>
<td>Win,X64,0,PS3,Android,IOS,5</td>
<td>C#,Visual Script</td>
<td>N</td>
<td>H</td>
<td>T</td>
<td>L</td>
</tr>
</tbody>
</table>

In contrast, the openness of the development process and scripts of Unity3D is easier for developers to use because customization and changes are not caused by the engine during the development process. Unity3D's rich resources provide a variety of scene models. AI plug-ins and rendering schemes are optimized, and their rendering configuration environment and physics simulation engine are more suitable for small team development.

### 3 Marker Matching and Detection

In the process of detecting and recognizing the plane marker, firstly, the connected video image is analyzed for the connected domain, and then the identifier matching is performed. In the process of marking the matching, the calculation of the transformation matrix is also required for each matched tag.

![Fig. 3.1 Marker Matching Process.](image)

1. Get the camera's image;
2. the identification of identification maps;
3. Location and direction detection of identification maps;
4. Combining virtual reality with the image on the camera
Graphic detection does not require complicated equipment and can achieve the required positioning accuracy. In mark-based graphic detection systems, the positions of various virtual objects are often marked by matching predefined planar graphic templates.

Thresholding is one of the most important methods of image segmentation. Assume that the grayscale histogram corresponds to an image \( f(x,y) \), which consists of dark objects on a bright background, and the pixels that make up the object and background constitute its two main modes. One way to extract objects from the background of an image is to select a threshold \( T \) to separate the two patterns. For any point \( (x, y) \), if its gray value \( f(x, y) \geq T \), the pixel point can be considered as a background point; otherwise it is considered as a marker point. A scaled image can be defined as

\[
g(x,y) = \begin{cases} 
1, & f(x) < T \\
0, & f(x) \geq T 
\end{cases}
\]  

(2)

Among them, the pixel marked as 1 belongs to the marker; and the pixel marked as 0 belongs to the background.

Fig. 3.2 Contrast the original image with the logo.

4 Design of an IoT System Based on Augmented Reality

The augmented reality system is divided into a basic environmental information display module, an environment rating alarm module, and a working hours reminder module according to function modules. The basic environmental information display module is mainly used to realistically present the basic conditions of the ambient temperature, dust, gas concentration, and Carbon Monoxide in the current environment, rendering these abstract environmental information in a concrete and real environment. This is the basic function of the system. Of course, only if these environmental information is of little significance, the environmental rating alarm module of the system will collect the environmental information data after passing the multi-sensor data fusion algorithm and then pass the rating to render the abstract result in the former environment. Provide miners with "visual" environmental information. The other is the working hours reminder module, calculated according to the 8-hour work schedule, removing the pre-shift and well cleaning time, calculated according to the 6-hour downhole working hours, and reminding the miners to rest every two hours.

To sum up, the functional framework diagram of the augmented reality part of this system is shown as follows:

According to the functional framework of this part, this article comprehensively uses a variety of augmented reality technologies to establish a man-machine-ring IoT augmented reality system. The key technology that I use is mainly real-time rendering technology. At the same time, the system is designed and implemented based on the interface provided by Unity3D.

5 Object Recognition Based on Augmented Reality for Internet of Things

This article designed the IoT Augmented Reality System. The previous chapters elaborated on the research contents and algorithms, realized the construction of the client of the AR system, the configuration of the server, and completed the design of the entire system. This chapter mainly deals with the construction of the system. The system is tested.

Fig. 5.1 Boot Interface.

Fig. 5.2 Data moves with camera movement.

Fig. 5.3 Augmented Reality Effects of Different Environment Levels.
6 Conclusion

This article is based on the existing underground IoT environment information collection system, capture the marker through the AR client. Really and steadily renders the processed environment information sent by the server, object recognition of environmental information in the Internet of things based on augmented reality system implemented.

In order to build a downhole AR system, Marker's three-dimensional tracking registration technology is used to enable the system to quickly detect the presence of Marker in a real environment, so that virtual information can be registered within the space of Marker's location.

The traditional smart mining IoT system uses sound and light alarms. Such alarming methods are obsolete and have poor visibility. In combination with the emerging augmented reality technology, an augmented reality system for the Internet of Things has been established. The virtual environment information is truly displayed in the real environment, which enhances the visualization of environmental information and greatly improves the miner's response to danger. Through comprehensive environmental testing, it is proved that this system is feasible.

The experimental results show that this system is workable. Object recognition of environmental information in the Internet of things based on augmented reality system designed and implemented in this paper, in providing a new direction for Research on IoT Smart Mines and Application of Mobile Devices.

Acknowledgments

This research was sponsored by the Natural Science Foundation of China (51704229), Key Scientific and Technological Project of Shaanxi Province (2016GY-040), and the Science Foundation of Xi’an University of Science and Technology (104-631990001). We also thank Master students Huan Li and Min Sun for their support of dataset collection.

References

5. Taçgün Z, Arslan A. The perceptions of CEIT postgraduate students regarding reality concepts: Augmented, virtual, mixed and mirror reality. Education & Information Technologies, 22,3(2016)